



Legacy of the 2008 Subprime Crisis

**Data Science in the
service of Financial
Technology**

Lecture plan

- Part 1 - The 2008 subprime crisis
 - The crisis in a nutshell
 - Securitization - missing information, obscure risks and moral hazard
 - The impact of the crisis and its legacy
 - Democratization of Finance
 - Data driven assessment of credit worthiness
 - Lending Club as a microcosms of credit risk
- Part 2 - Data Structures for Data Science
 - Python numerical library - NumPy
 - Python financial data library - Pandas

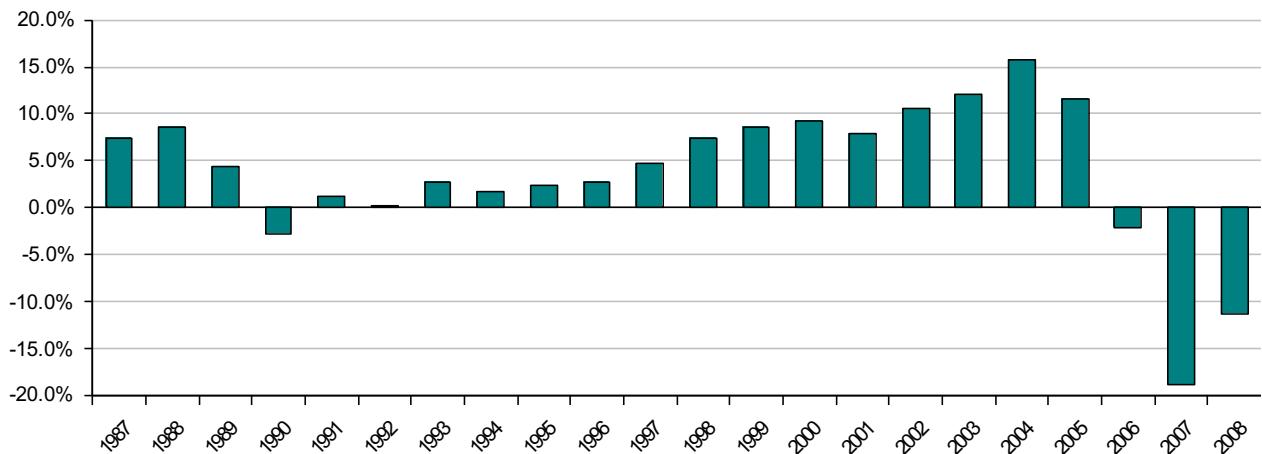
The 2008 Subprime Crisis

Timeline of the Crisis

Feb 2007	DJ at record 12,786
Apr 2007	New Centry files chapter 11
Aug 2007	BNP blocks withdrawal from two funds
Aug 2007	Fed takes down rate
Dec 2007	Fed offers short term credit to banks holding subprime mortgages
Feb 2008	Economic Stimulus Act of 2008
Mar 2008	Bear Stern bailout
Jul 2008	IndyMac fails
Jul 2008	Housing and Economic Recovery Act
Sep 7, 2008	Fannie Mae and Freddie Mac taken over by federal government
Sep 15, 2008	DJ drops 504 points in a day
Sep 15, 2008	Lehman Brothers file chap. 11
Sep 16, 2008	Fed takes over AIG
Sep 21, 2008	Goldman Sachs and Morgan Stanley convert to 'bank holding' status
Sep 26, 2008	Washington mutual goes under after run on the bank
Sep 29, 2008	House rejects Troubled Asset Relief program (TARP)
Oct 2008	Economic Stabilization Act
Nov 2008	Asset backed loan facility
Dec 2008	Fed lowers rate to zero
Jan 2009	GM, Ford and Chrysler receive bailout under TARP
Feb 2009	Stimulus package approved by congress
Mar 2009	DJ hits lowest level 6,443

House Price Change

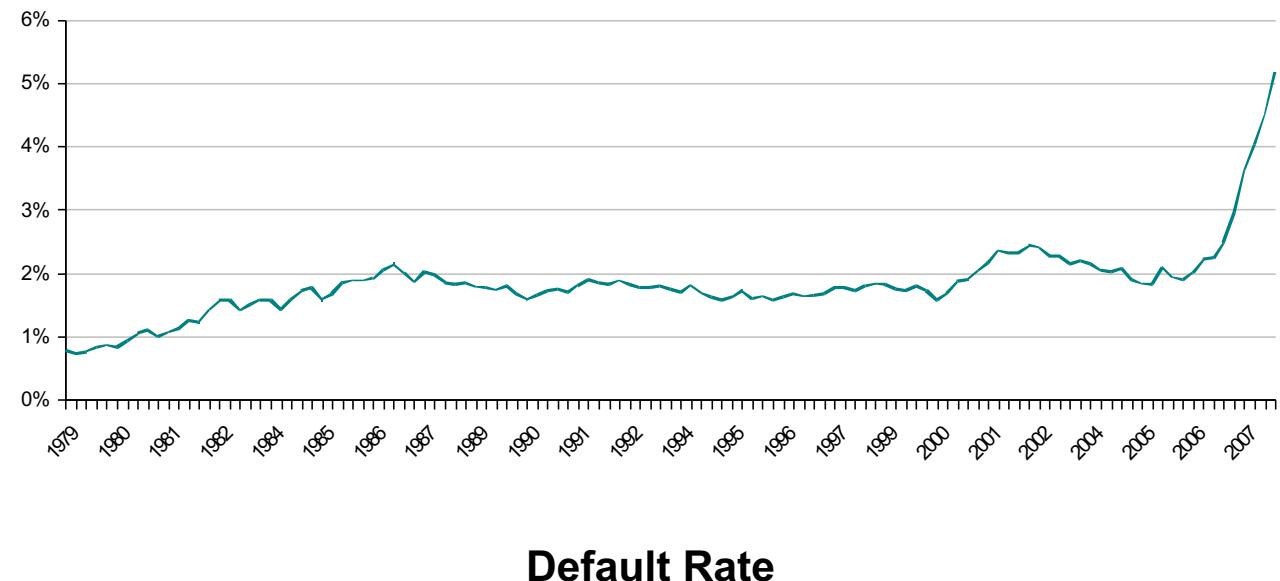
- Housing prices were relatively stable during the 1990s, but they began to rise toward the end of the decade.
- Between January 2002 and mid-year 2006, housing prices increased by a whopping 87 percent.
- The boom had turned to a bust, and the housing price declines continued throughout 2007 and 2008.
- By the third quarter of 2008, housing prices were approximately 25 percent below their 2006 peak.



Annual Existing House Price Change

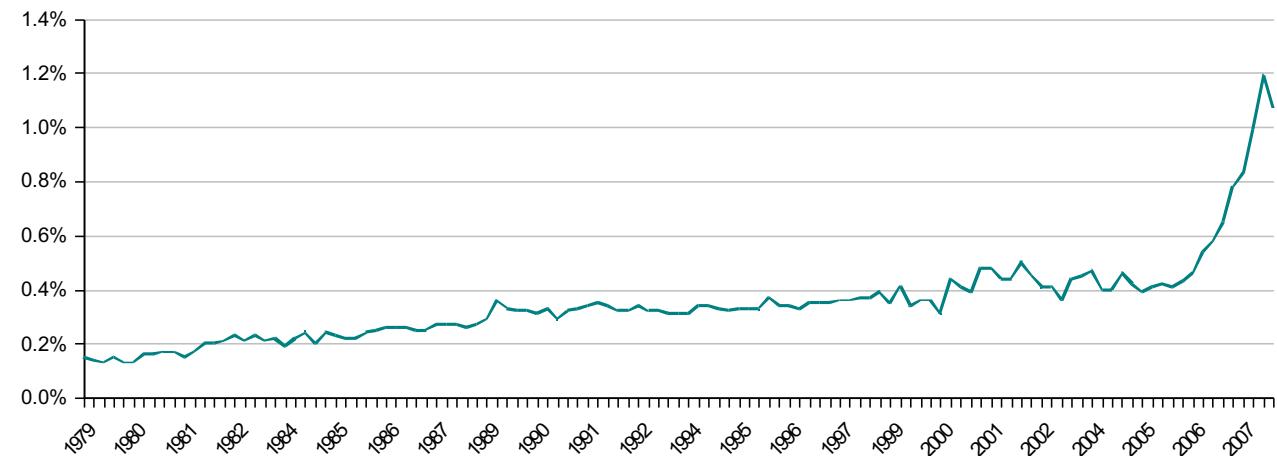
The Default Rate

- The default rate fluctuated, within a narrow range, around 2 percent prior to 2006.
- It increased only slightly during the recessions of 1982, 1990, and 2001.
- The rate began increasing sharply during the second half of 2006
- It reached 5.2 percent during the third quarter of 2008.



Foreclosure Rate

- Housing prices were relatively stable during the 1990s, but they began to rise toward the end of the decade.
- Between January 2002 and mid-year 2006, housing prices increased by a whopping 87 percent.
- The boom had turned to a bust, and the housing price declines continued throughout 2007 and 2008.
- By the third quarter of 2008, housing prices were approximately 25 percent below their 2006 peak.



Foreclosure Rate

Major Drivers of the Crisis



Home ownership in the US



Monetary Policy

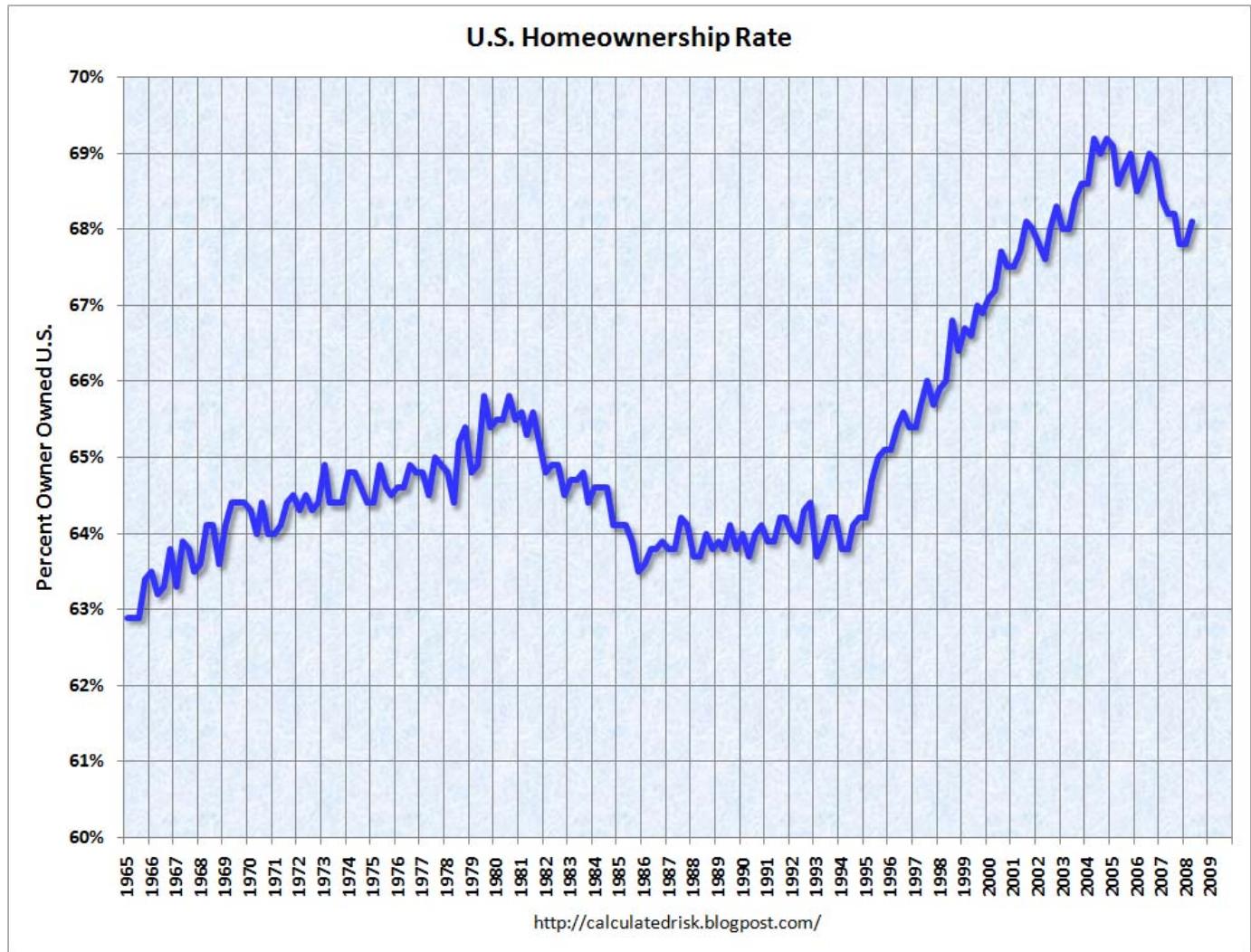


Securitization



Speculation

Home Ownership in the US



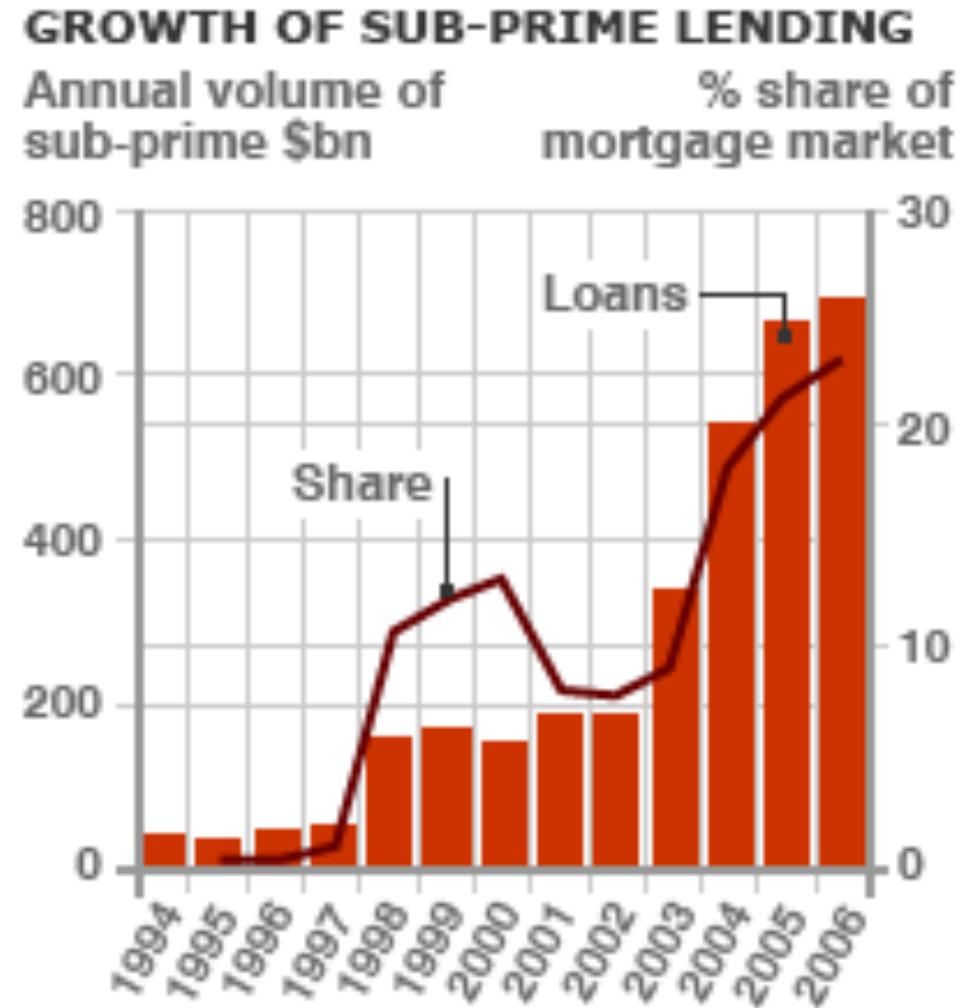
Home Ownership in the US

Beginning in the mid-1990s, government regulations began to erode the conventional lending standards.

- Fannie Mae and Freddie Mac hold a huge share of American mortgages.
- Beginning in 1995, **HUD regulations** required Fannie Mae and Freddie Mac to increase their holdings of loans to low and moderate income borrowers.
- **HUD regulations** imposed in 1999 required Fannie and Freddie to accept more loans with little or no down payment.
- 1995 regulations stemming from an extension of the **Community Reinvestment Act** required banks to extend loans in proportion to the share of minority population in their market area. Conventional lending standards were reduced to meet these goals.

Home Ownership in the US

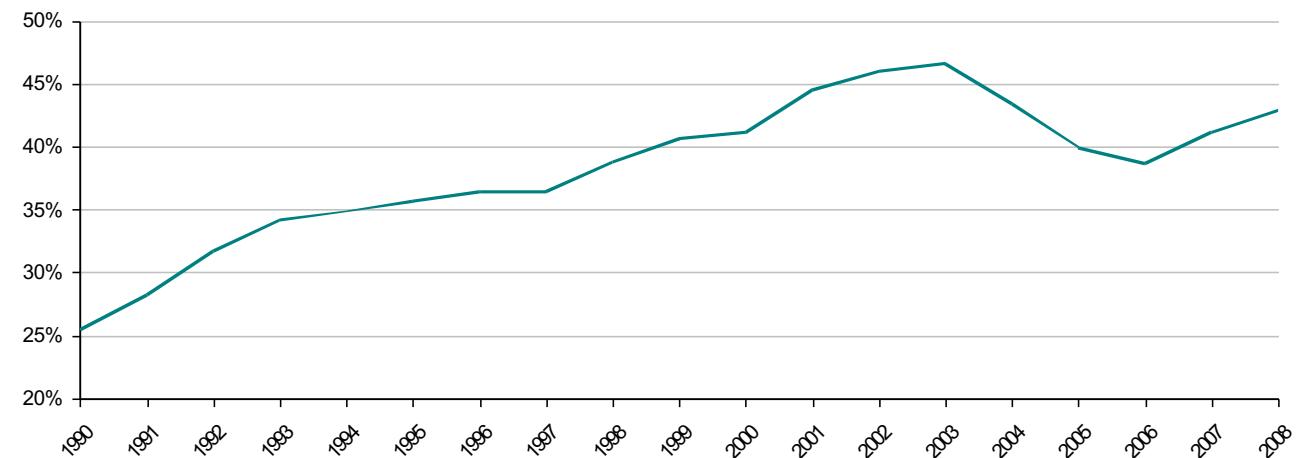
Subprime refers to loans that do not meet *Fannie Mae* or *Freddie Mac* guidelines due to credit status, income and job history, income to mortgage payment ratio, etc.



SOURCE: Center for Responsible Lending /Inside Mortgage Finance

Fannie Mae/Freddie Mac Share

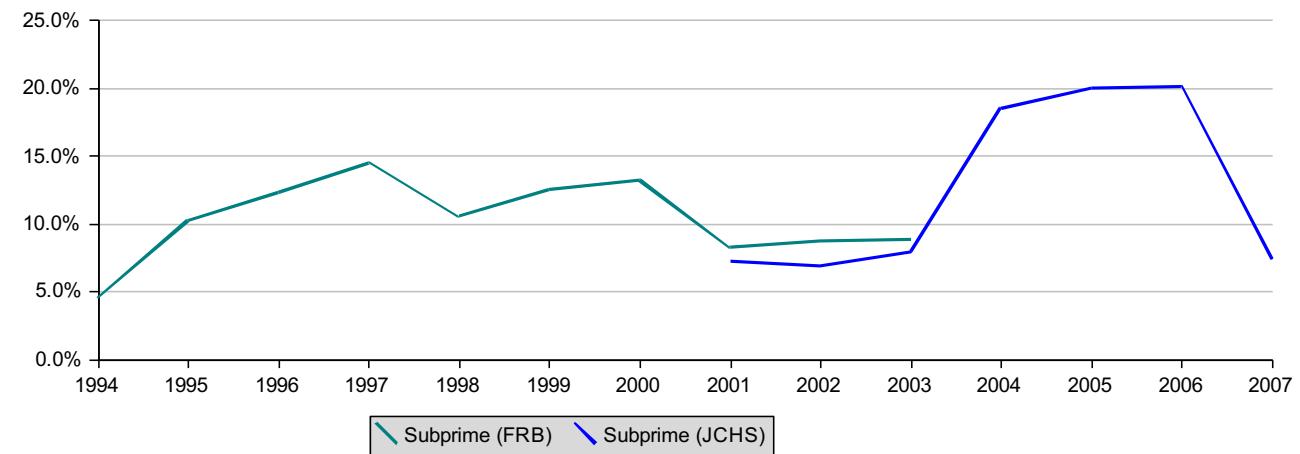
- The share of all mortgages held by Fannie Mae and Freddie Mac rose from 25 percent in 1990 to 45 percent in 2001.
- Their share has fluctuated modestly around 45 percent since 2001.



Freddie Mac/Fannie Mae Share of Outstanding Mortgages

Fannie Mae/Freddie Mac Share

- Subprime mortgages as a share of total mortgages originated during the year, increased from 5% in 1994 to 13% in 2000 and on to 20% in 2004-2006.



Subprime Mortgage Originations as a Share of Total

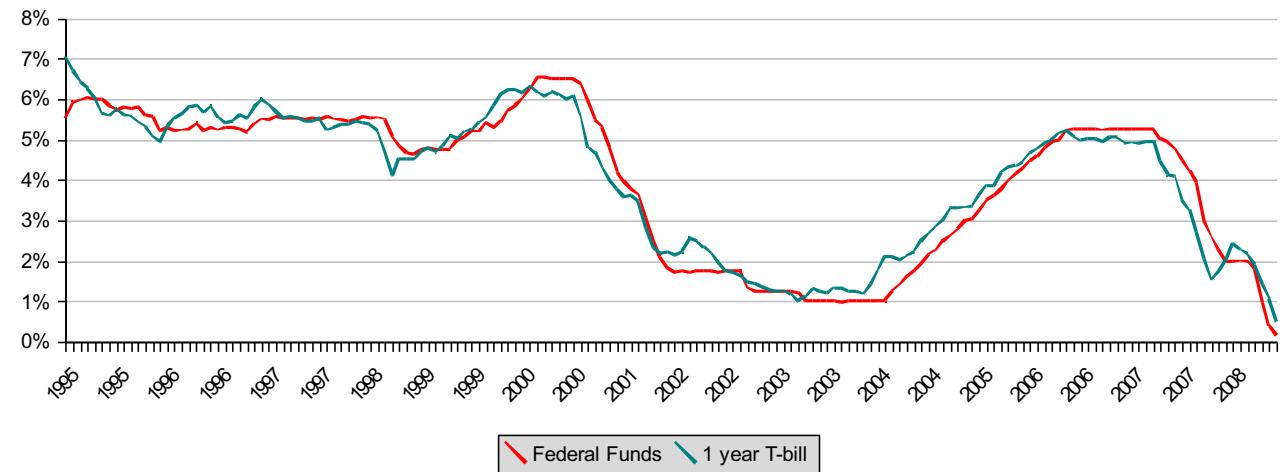
Monetary Policy

The Fed's interest rates policy during 2002-2006

- Fed's prolonged Low-Interest Rate Policy of 2002-2004 increased demand for, and price of, housing
- The low short-term interest rates made adjustable rate loans with low down payments highly attractive
- As the Fed pushed short-term interest rates upward in 2005-2006, adjustable rates were soon reset, monthly payment on these loans increased, housing prices began to fall, and defaults soared

Fannie Mae/Freddie Mac Share

- The Fed injected additional reserves and kept short-term interest rates at 2% or less throughout 2002-2004.
- Due to rising inflation in 2005, the Fed pushed interest rates upward.
- Interest rates on adjustable rate mortgages rose and the default rate began to increase rapidly



Federal Funds Rate and 1-Year T-Bill Rate

Securitization

An SEC Rule change adopted in April 2004 led to highly leverage lending practices by investment banks and their quick demise when default rates increased

- The rule favoured lending for residential housing
- Loans for residential housing could be leveraged by as much as 25 to 1, and as much as 60 to 1, when bundled together and financed with securities
- Based on historical default rates, mortgage loans for residential housing were thought to be safe. But this was no longer true because regulations had seriously eroded the lending standards and the low interest rates of 2002-2004 had increased the share of ARM loans with little or no down payment
- When default rates increased in 2006 and 2007, the highly leveraged investment banks soon collapsed

MBS & CDO

- A **mortgage-backed security (MBS)** is a “bond” whose cash flows are backed by the principal and interest payments of a set of mortgage loans.
- **Collateralized debt obligations (CDOs)** are an unregulated type of asset-backed security and structured credit product. CDOs are constructed from a portfolio of fixed-income assets. These assets are divided by the ratings firms that assess their value into different tranches.

Figure 1 - Simple, typical CDO Tranche Structure

Tranche	Percent of capital structure	Rating	Coupon
Class A	77.50	AAA	LIBOR + 26
Class B	9.00	A	LIBOR + 75
Class C	2.75	BBB	LIBOR + 180
Class D	2.75	BB	LIBOR + 475
Preferred shares	8.00	NR	Residual cash flow

Credit Default Obligations

A Primer on Credit Default Insurance

THE PLAYERS

The insurance buyer is often a bond investor seeking protection against default on an asset he owns. But many are speculators, who do not own the asset, but use the credit default swap to bet on the health of a company.



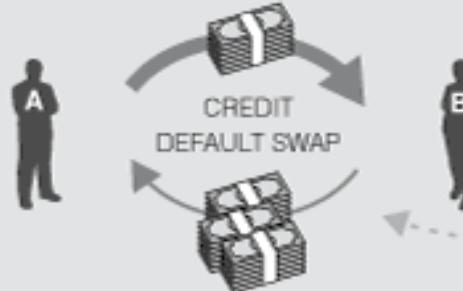
The insurance seller is often a hedge fund, insurance company or bank. It receives premiums from the buyer throughout the contract.



THE PROCESS

1

Party A **buys credit default insurance** from Party B to protect against default on a bond, or to bet on a company's health.



2

In the case of a default, Party B would pay the bond's full value to Party A.

THE PROBLEM

3

Party B can **assign the insurance contract to another party**

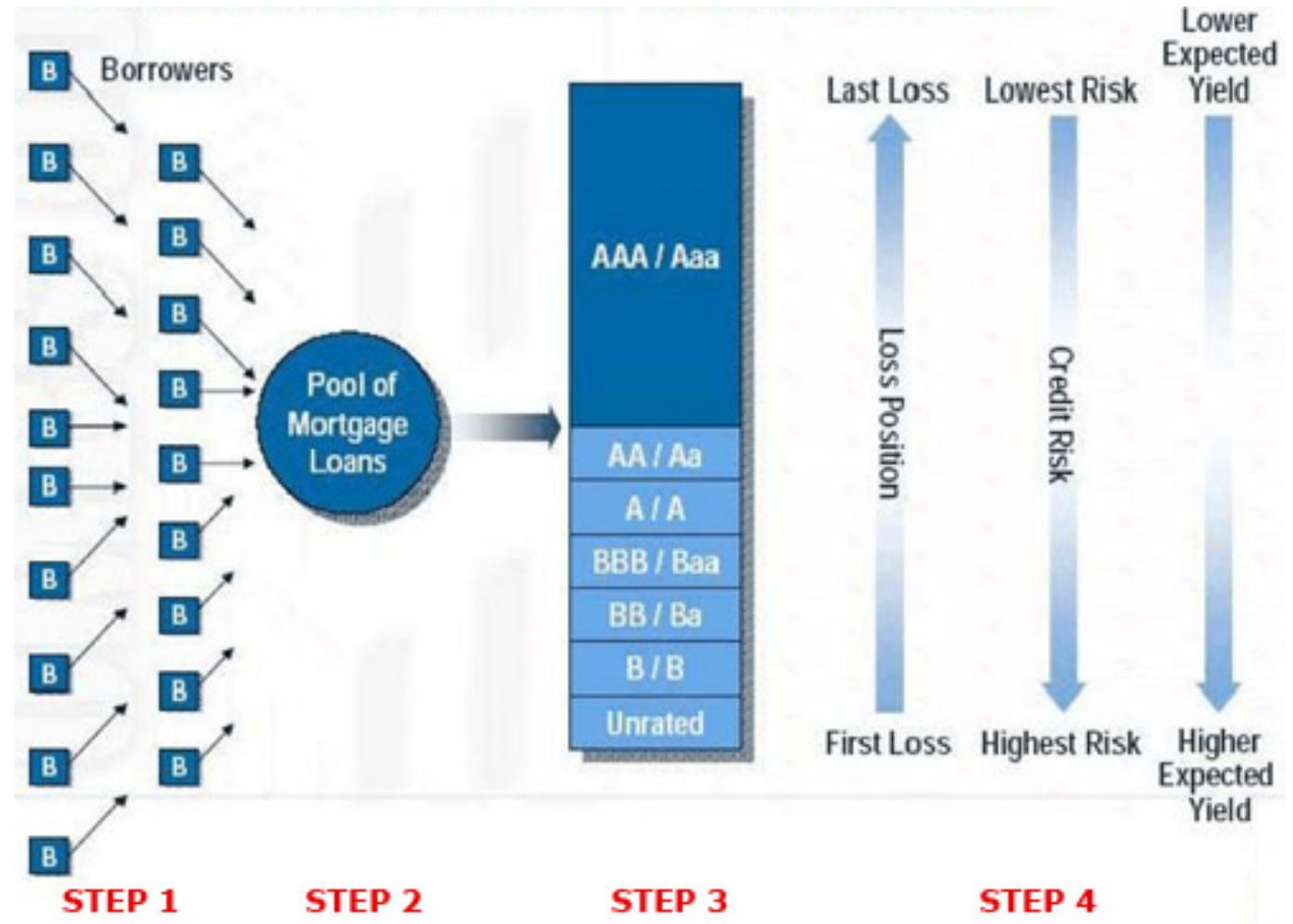
...who can assign it to another...

...and he to another.



In the case of a default, Party A may have to track down the final party in the insurance agreement. However, this party **may or may not be in a position to pay the bond's full value.**

CDOs trading on asymmetric information

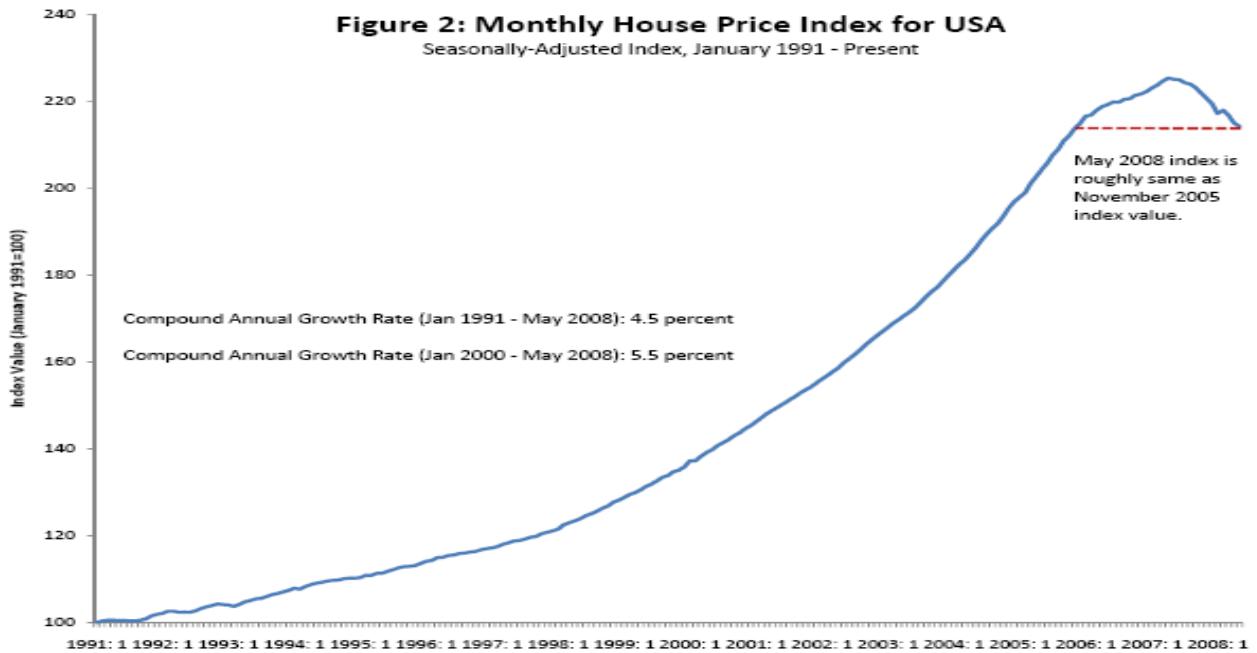


Speculation

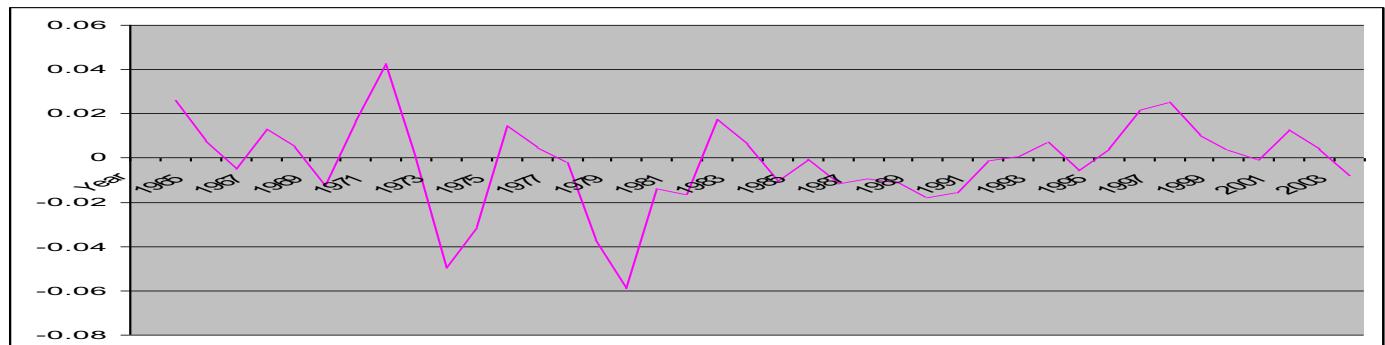


House Price

Source: Office of Federal Housing Enterprise Oversight

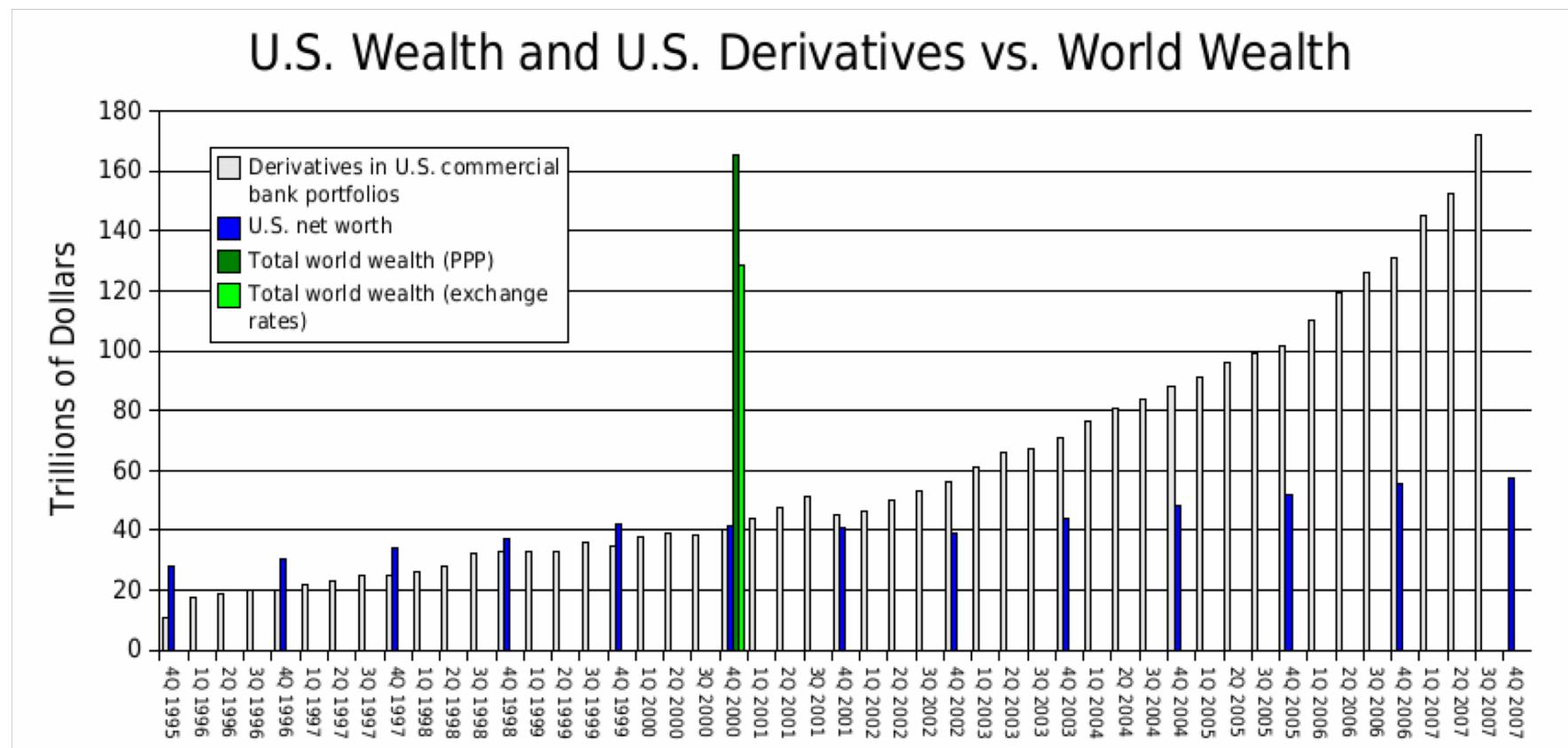


House Prices



Real Salary

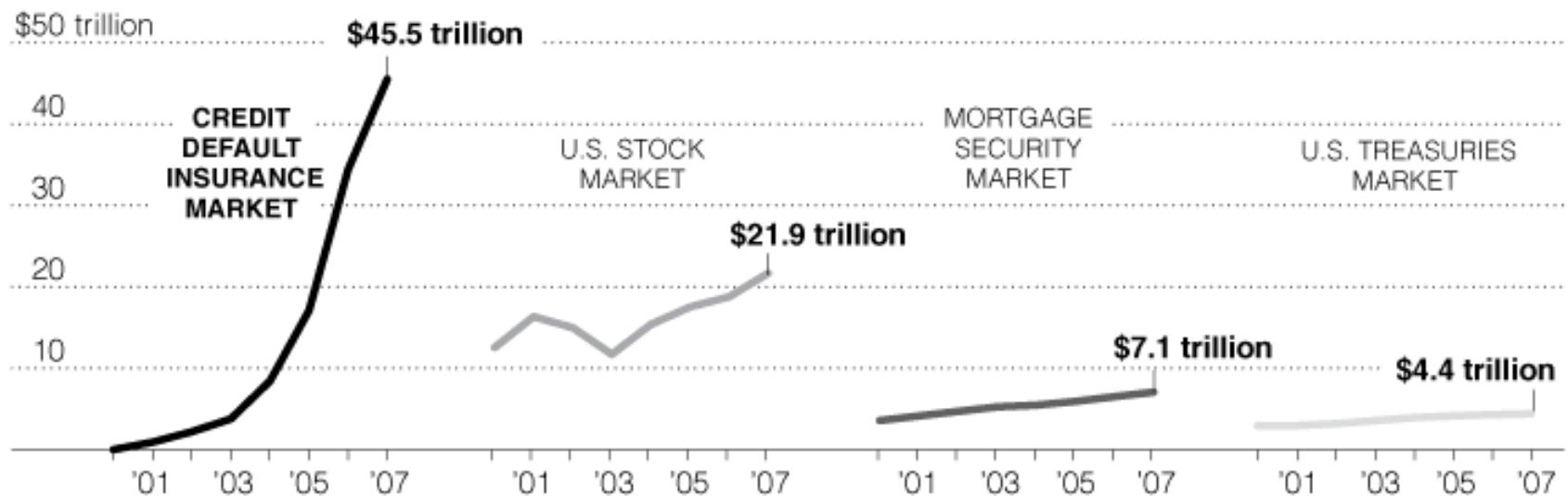
Speculative Fever (1)



Speculative Fever (2)

In the Shadow of an Unregulated Market

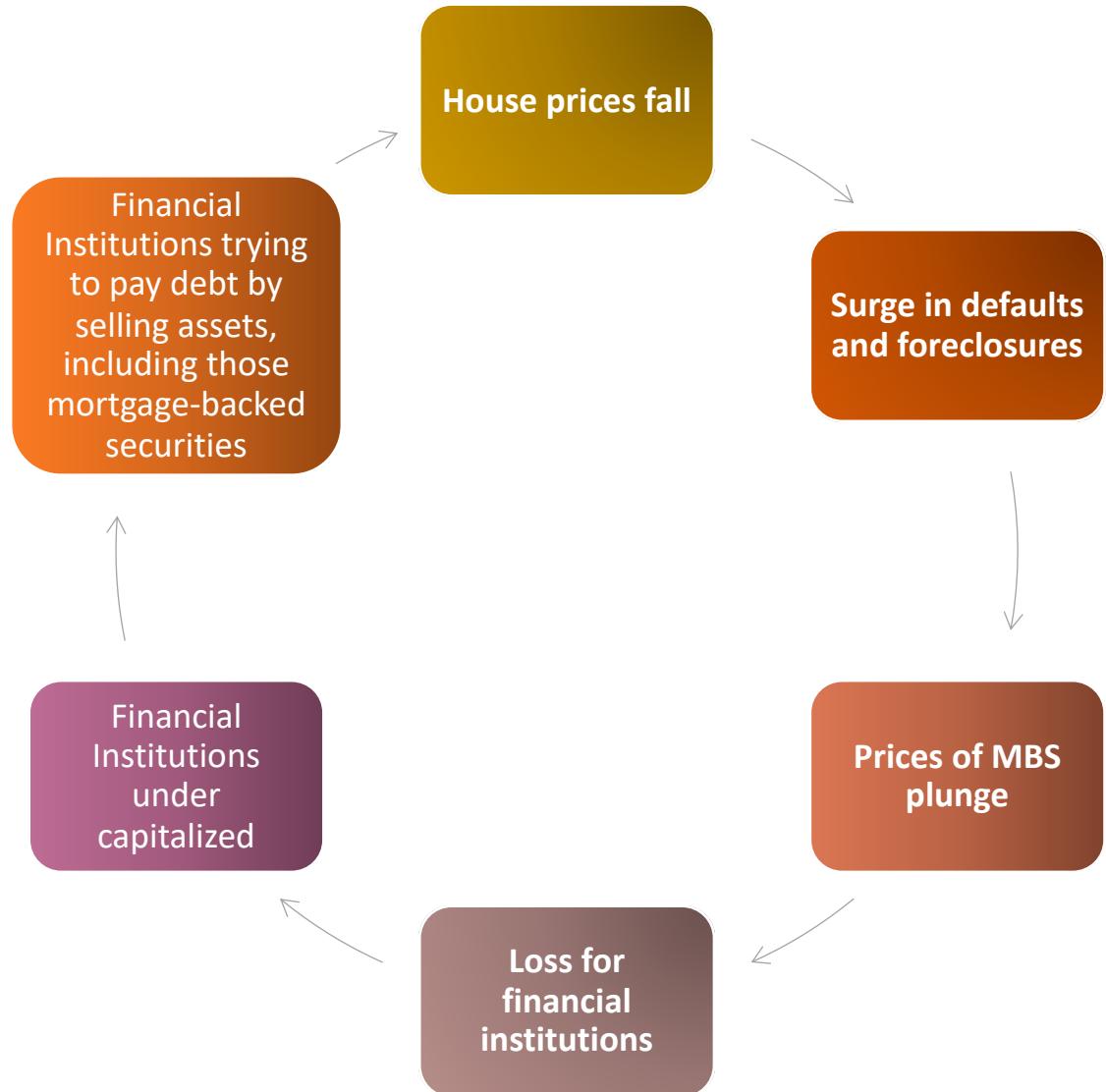
The value of the credit default insurance market is now much larger than the domestic stock market, mortgage securities market and United States Treasuries market.



Sources: Thomson Proprietary Research; International Swaps and Derivatives Association

THE NEW YORK TIMES

The Vicious Circle



Government Bailouts

Bear Stearns (\$29 bn loan guarantee)

AIG (\$86 bn bridge loan)

Fannie Mae and Freddie Mac (\$200 bn pledges of protection)

Through June 2008, the Fed had provided approximately \$1.2 trillion in loans to various financial institutions through its Term auction facility.

\$700 bn (5% of US GDP) to purchase large amounts of illiquid, risky mortgage backed securities from financial institutions

July 2010 Dodd–Frank Wall Street Reform and Consumer Protection Act

Timeline of the Crisis

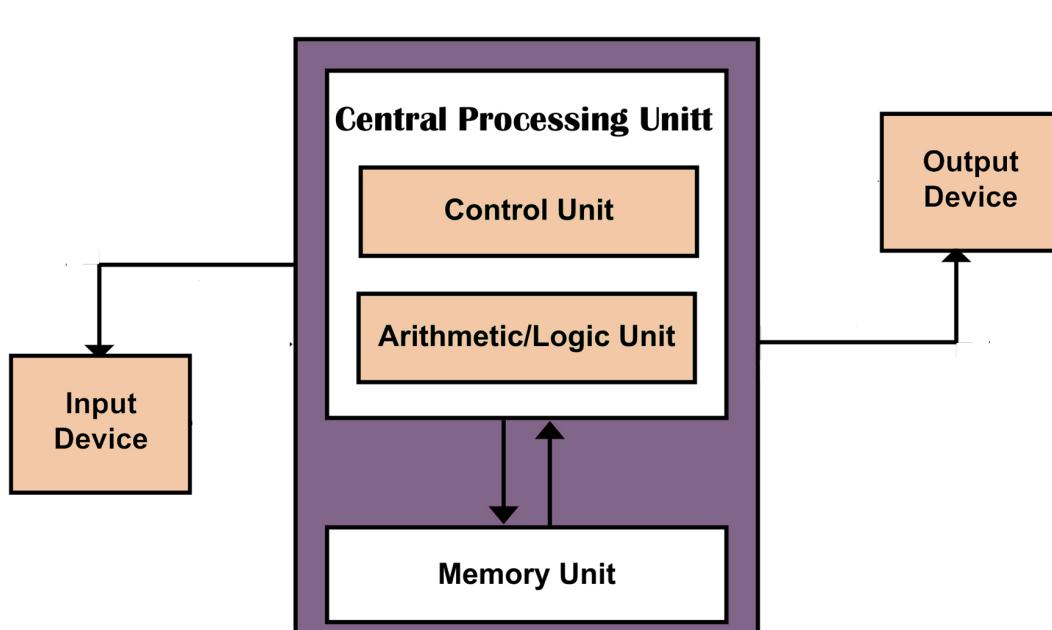
Feb 2007	DJ at record 12,786
Apr 2007	New Centry files chapter 11
Aug 2007	BNP blocks withdrawal from two funds
Aug 2007	Fed takes down rate
Dec 2007	Fed offers short term credit to banks holding subprime mortgages
Feb 2008	Economic Stimulus Act of 2008
Mar 2008	Bear Stern bailout
Jul 2008	IndyMac fails
Jul 2008	Housing and Economic Recovery Act
Sep 7, 2008	Fannie Mae and Freddie Mac taken over by federal government
Sep 15, 2008	DJ drops 504 points in a day
Sep 15, 2008	Lehman Brothers file chap. 11
Sep 16, 2008	Fed takes over AIG
Sep 21, 2008	Goldman Sachs and Morgan Stanley convert to 'bank holding' status
Sep 26, 2008	Washington mutual goes under after run on the bank
Sep 29, 2008	House rejects Troubled Asset Relief program (TARP)
Oct 2008	Economic Stabilization Act
Nov 2008	Asset backed loan facility
Dec 2008	Fed lowers rate to zero
Jan 2009	GM, Ford and Chrysler receive bailout under TARP
Feb 2009	Stimulus package approved by congress
Mar 2009	DJ hits lowest level 6,443



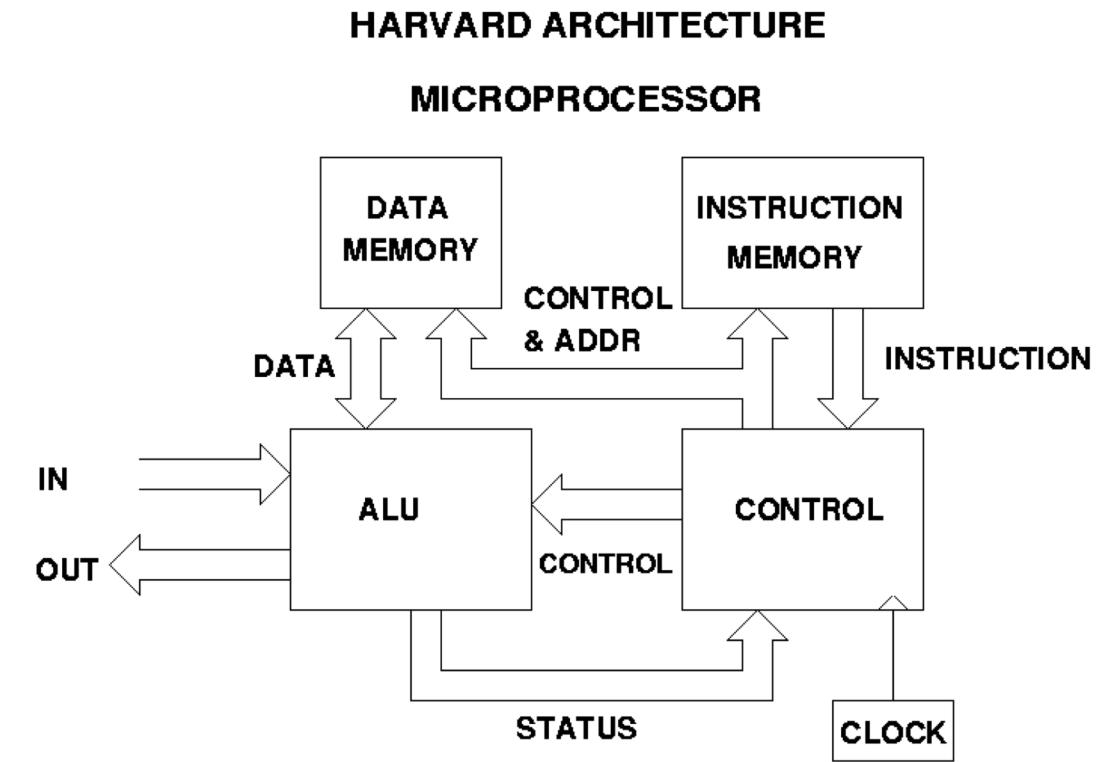
Python Data Structures for Data Science

- NumPy
- Pandas

Computer Architecture

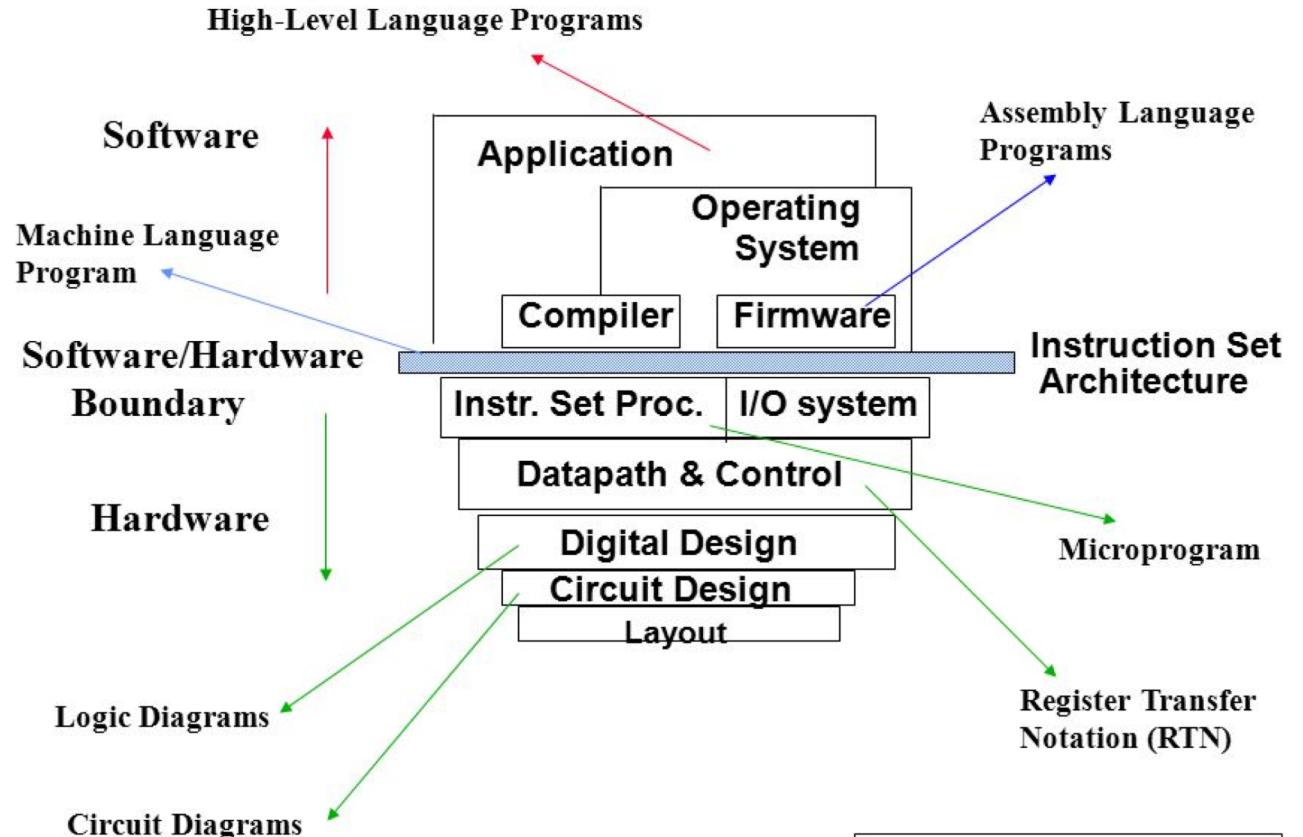


Von Neumann Architecture

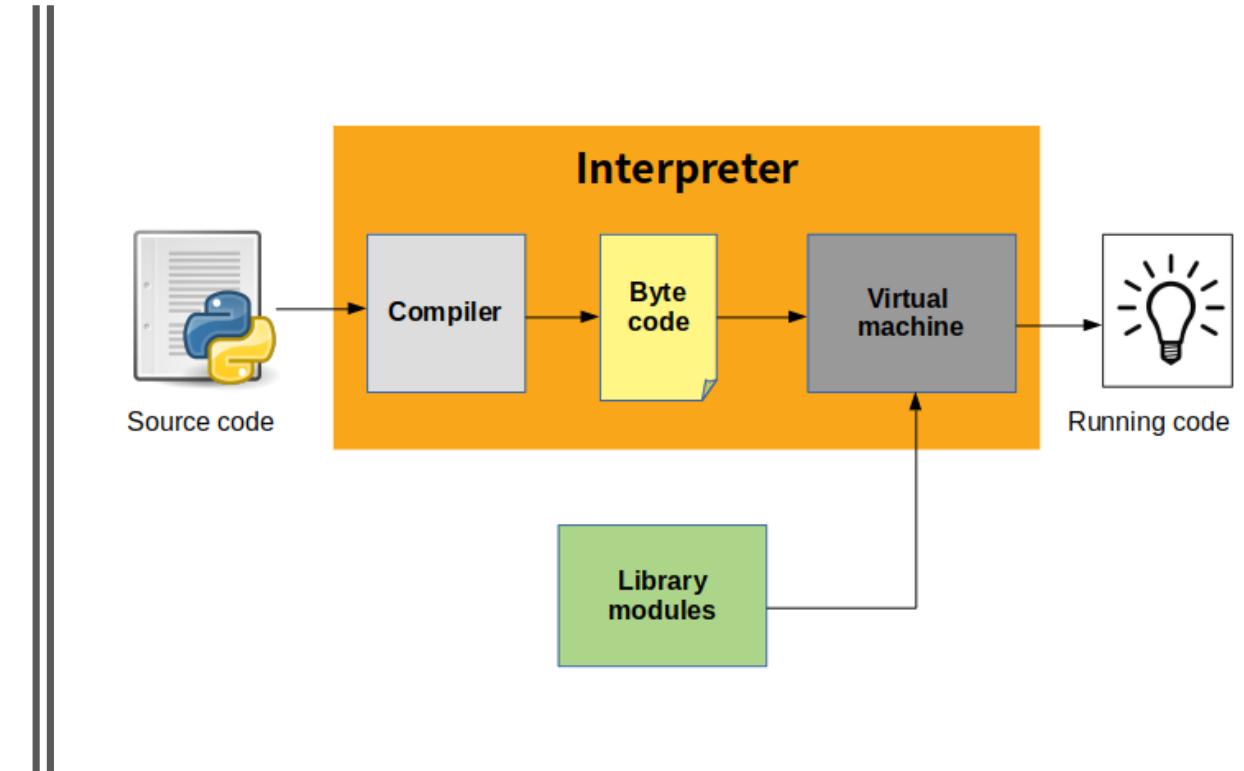
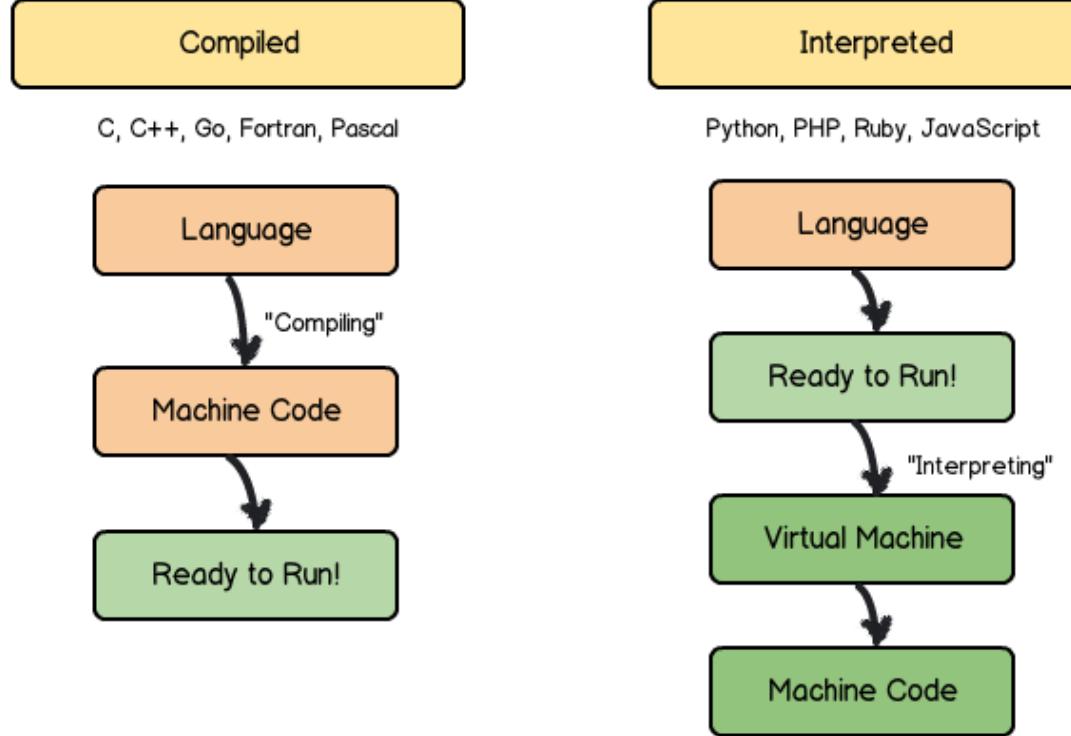


Computer Architecture

Hierarchy of Computer Architecture

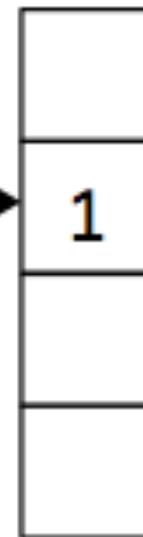


Interpreters



Dynamic vs
Static typed
languages

C Integer



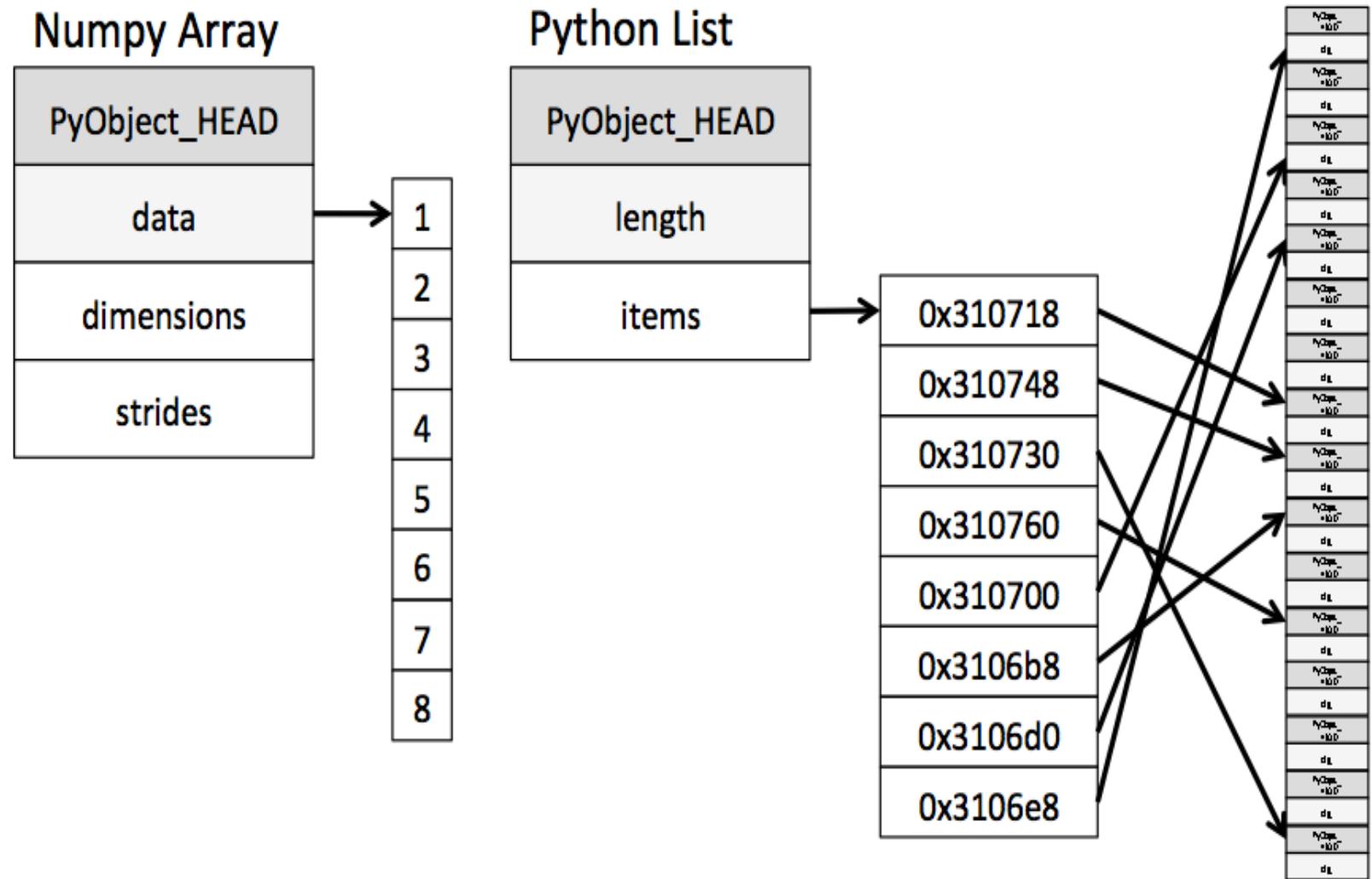
Python Integer

PyObject_HEAD

digit

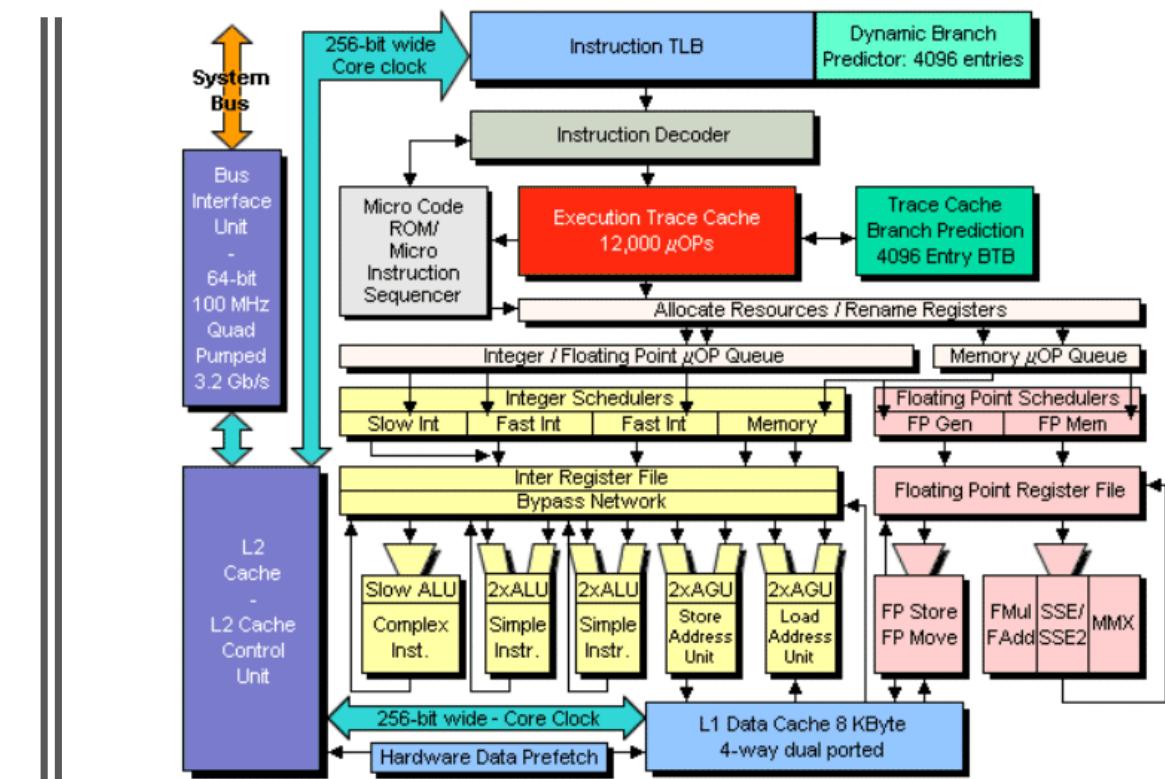
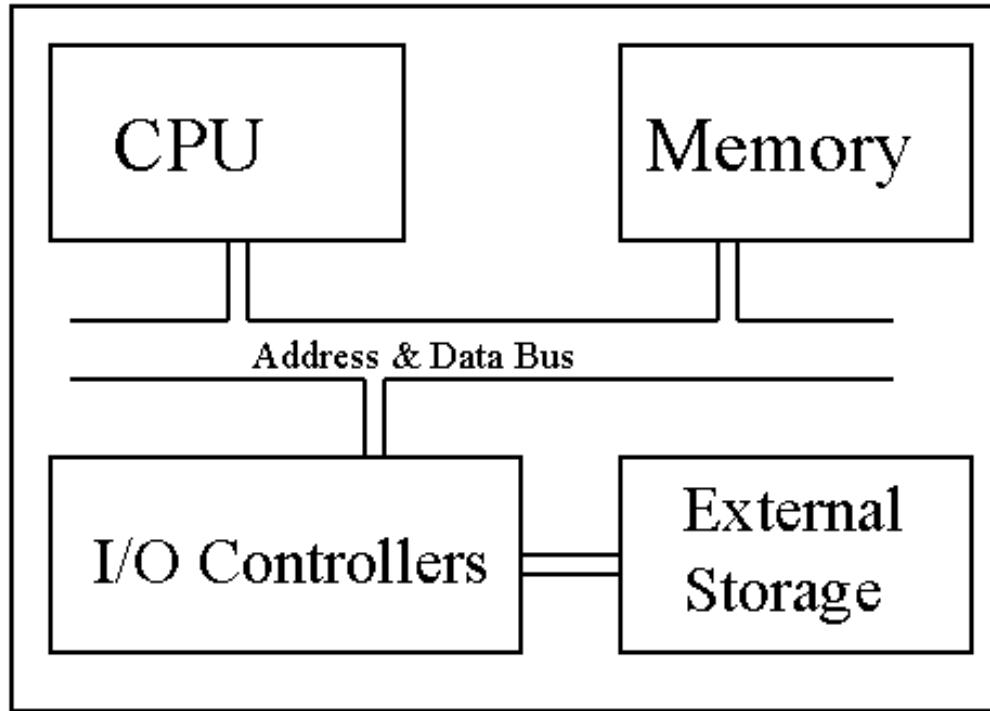
1

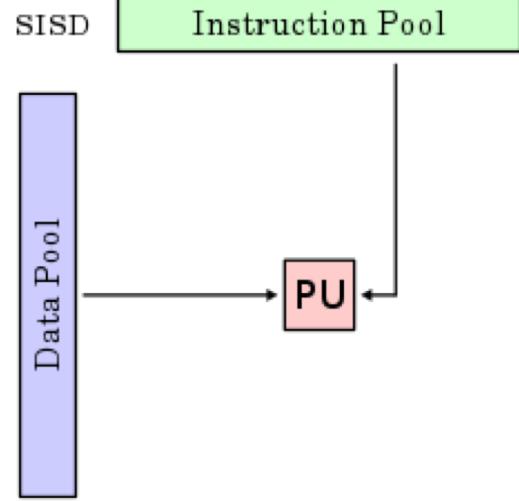
Dynamic vs Static typing



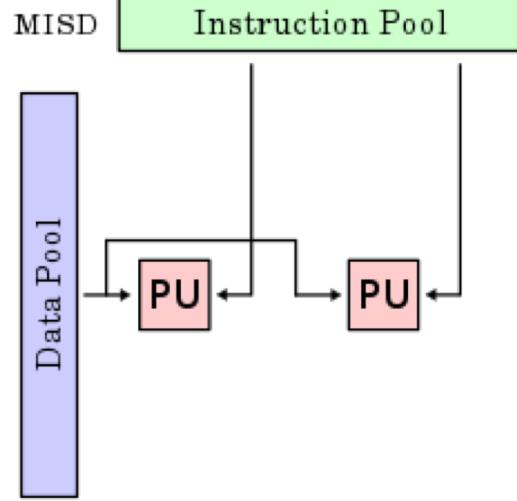
Computer Architecture

Basic Digital Computer Architecture

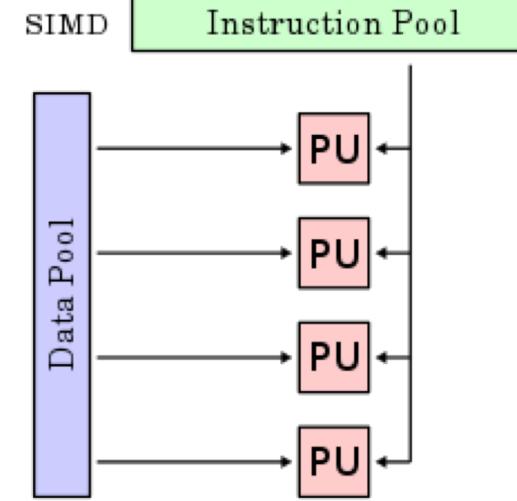




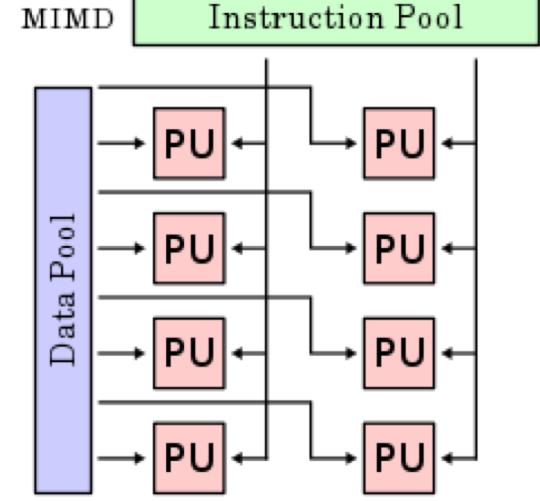
**Single instruction stream
single data stream (SISD)**



**Multiple instruction
streams, single data
stream (MISD)**



**Single instruction stream,
multiple data streams
(SIMD)**

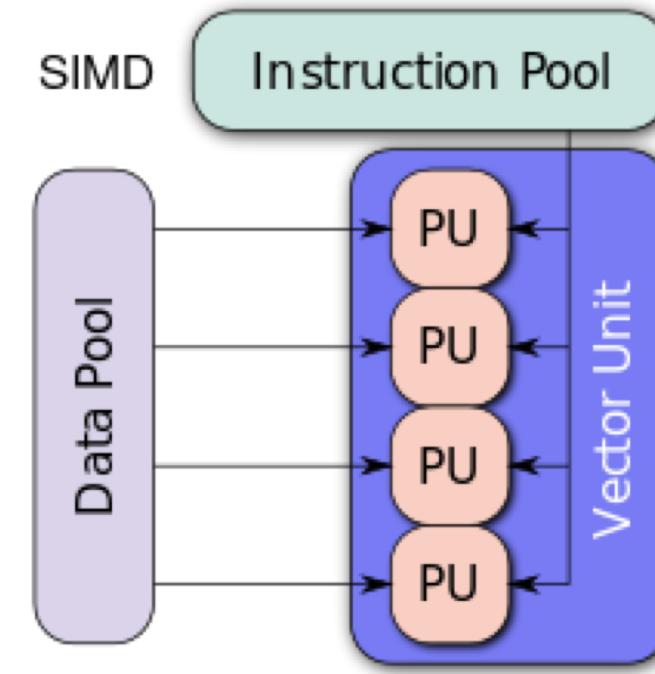


**Multiple instruction
streams, multiple data
streams (MIMD)**

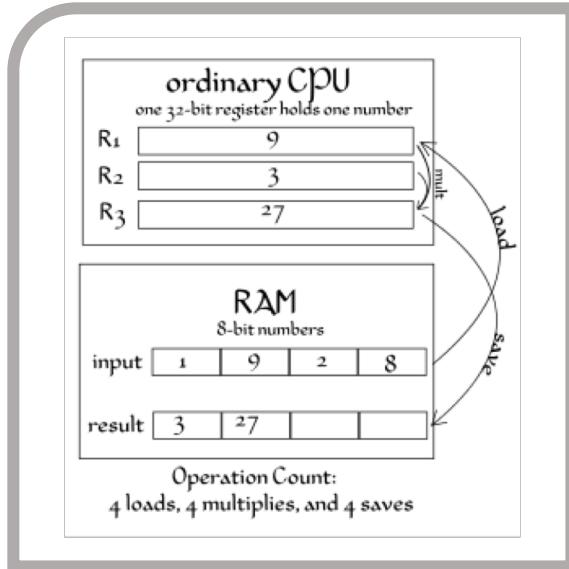
Microprocessor Architecture

Flynn's taxonomy

Vectorization

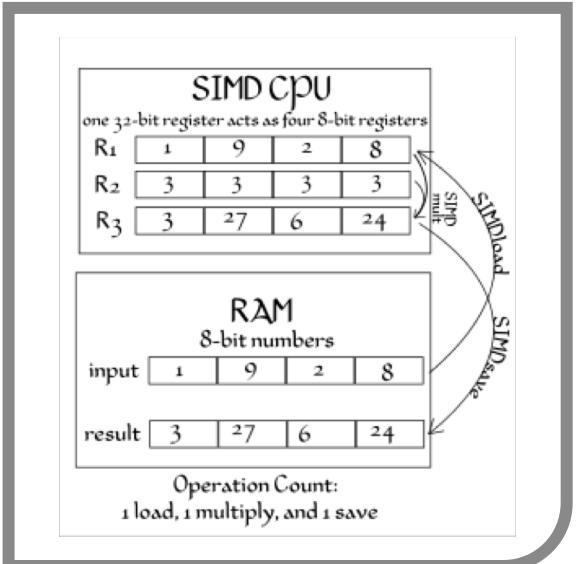


Vectorization

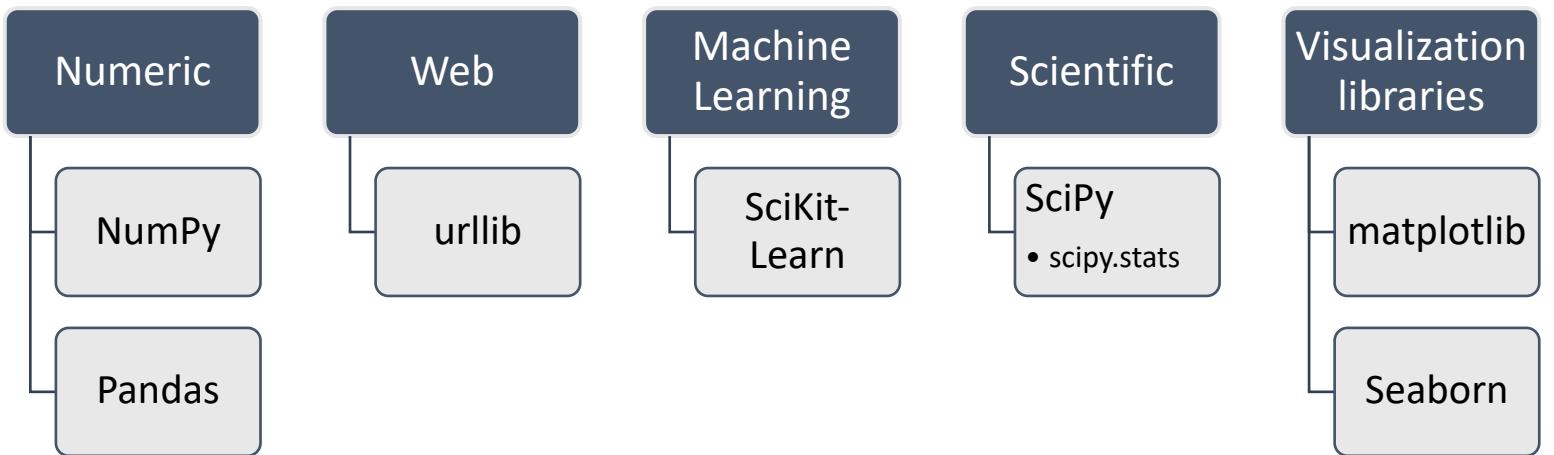


The ordinary tripling of four 8-bit numbers. The CPU loads one 8-bit number into R1, multiplies it with R2, and then saves the answer from R3 back to RAM. This process is repeated for each number.

The SIMD tripling of four 8-bit numbers. The CPU loads 4 numbers at once, multiplies them all in one SIMD-multiplication, and saves them all at once back to RAM. In theory, the speed can be multiplied by 4.



Python Libraries for Data Science



NumPy

- introduces objects for multidimensional arrays and matrices, as well as functions that allow to easily perform advanced mathematical and statistical operations on those objects
- provides ***vectorization*** of mathematical operations on arrays and matrices which significantly improves the performance
- many other python libraries are built on NumPy

Link: <http://www.numpy.org/>

Pandas

- adds data structures and tools designed to work with table-like data (similar to Series and Data Frames in R)
- provides tools for data manipulation: reshaping, merging, sorting, slicing, aggregation etc.
- allows handling missing data

Link: <http://pandas.pydata.org/>

SciPy

- collection of algorithms for linear algebra, differential equations, numerical integration, optimization, statistics and more
- part of SciPy Stack
- built on NumPy

Link: <http://www.numpy.org/>

SciKit-Learn

- provides machine learning algorithms: classification, regression, clustering, model validation etc.
- built on NumPy, SciPy and matplotlib

Link: <http://scikit-learn.org/>

matplotlib

- python 2D plotting library which produces publication quality figures in a variety of hardcopy formats
- a set of functionalities similar to those of MATLAB
- line plots, scatter plots, barcharts, histograms, pie charts etc.
- relatively low-level; some effort needed to create advanced visualization

Link: <https://matplotlib.org/>

Seaborn

- based on matplotlib
- provides high level interface for drawing attractive statistical graphics
- Similar (in style) to the popular ggplot2 library in R

Link: <https://seaborn.pydata.org/>

Data Frame Types

Pandas Type	Native Python Type	Description
object	string	The most general dtype. Will be assigned to your column if column has mixed types (numbers and strings).
int64	int	Numeric characters. 64 refers to the memory allocated to hold this character.
float64	float	Numeric characters with decimals. If a column contains numbers and NaNs(see below), pandas will default to float64, in case your missing value has a decimal.
datetime64, timedelta[ns]	N/A (but see the datetime module in Python's standard library)	Values meant to hold time data. Look into these for time series experiments.

Data Frame Attributes

df.attribute	description
dtypes	list the types of the columns
columns	list the column names
axes	list the row labels and column names
ndim	number of dimensions
size	number of elements
shape	return a tuple representing the dimensionality
values	numpy representation of the data

Data Frame Methods

df.method()	description
head([n]), tail([n])	first/last n rows
describe()	generate descriptive statistics (for numeric columns only)
max(), min()	return max/min values for all numeric columns
mean(), median()	return mean/median values for all numeric columns
std()	standard deviation
sample([n])	returns a random sample of the data frame
dropna()	drop all the records with missing values

Data Frame Missing Data

df.method()	description
dropna()	Drop missing observations
dropna(how='all')	Drop observations where all cells is NA
dropna(axis=1, how='all')	Drop column if all the values are missing
dropna(thresh = 5)	Drop rows that contain less than 5 non-missing values
fillna(0)	Replace missing values with zeros
isnull()	returns True if the value is missing
notnull()	Returns True for non-missing values

Basic Descriptive Statistics

df.method()	description
describe	Basic statistics (count, mean, std, min, quantiles, max)
min, max	Minimum and maximum values
mean, median, mode	Arithmetic average, median and mode
var, std	Variance and standard deviation
sem	Standard error of mean
skew	Sample skewness
kurt	kurtosis

Graphics

description	
distplot	histogram
barplot	estimate of central tendency for a numeric variable
violinplot	similar to boxplot, also shows the probability density of the data
jointplot	Scatterplot
regplot	Regression plot
pairplot	Pairplot
boxplot	boxplot
swarmplot	categorical scatterplot
factorplot	General categorical plot